Computer Physics Communications 179 (2008) 716–723

Contents lists available at ScienceDirect

Computer Physics Communications

www.elsevier.com/locate/cpc

An objective, model-independent method for detection of non-uniform steps in noisy signals

Bennett Kalafut, Koen Visscher *

Department of Physics, University of Arizona, Tucson, AZ 85721, USA

ARTICLE INFO

Article history:
Received 15 February 2008
Received in revised form 11 June 2008
Accepted 12 June 2008
Available online 20 June 2008

PACS:
87.80.Nj
02.50.Tt

Keywords:
Step detection
Change point analysis
Single molecule biophysics

ABSTRACT

Biophysical techniques, such as single molecule FRET, fluorescence microscopy, single ion-channel patch clamping, and optical tweezers often yield data that are noisy time series containing discrete steps. Here we present a method enabling objective identification of nonuniform steps present in such noisy data. Our method does not require the assumption of any underlying kinetic or state models and is thus particularly useful for analysis of novel and poorly understood systems. In contrast to other model-independent methods, no parameters or other information is taken from the user. We find that, at high noise levels, our method exceeds the performance of other model-independent methods in accurately locating steps in simulated noisy data.

1. Introduction

Physical measurements of single (bio)molecules have enabled the detection of biochemical and structural states and the transitions between them, which otherwise would have been obscured in ensemble-averaged experiments. The time evolution of many such biomolecular systems as observed with single-molecule techniques may be characterized by dwells punctuated by rapid transitions, often called steps. A classic example results from the patch-clamp technique monitoring the open and closed states of a single trans-membrane channel [1]. More recent advances using optical tweezers and fluorescence microscopy have uncovered staircase displacement records of individual motor proteins [2–4], while single-pair FRET has provided unprecedented insight into structural heterogeneity and conformational kinetics of individual enzymes and ribozymes [5,6], to name only a few of an increasing arsenal of single-molecule methods.

Measuring the distributions of step sizes and durations, rather than just their mean value, has provided invaluable clues towards understanding the structural and kinetic mechanisms underlying biomolecules' function [7]. Thus it is imperative to be able to reliably detect the size and durations of steps in these data, which are typically noisy. When the size of the step far exceeds the typical noise levels associated with the experiment, one may choose to identify steps by eye, possibly without much danger of experimentalists' bias and without much penalty towards accuracy [8]. Such visual identification may be aided through use of filters such as low-pass filters, Chung and Kennedy's nonlinear filter [9], or Haar wavelet-based filters [10].

However, once S/N, the ratio of the step size and the standard deviation of the noise, decreases, such subjective methods no longer suffice and objective methods guaranteeing a high degree of reproducibility are required. Depending on the biomolecular system and the design of the experiment, these methods may be as simple as constructing a histogram of the raw data or computation of correlation functions to uncover step sizes and rates, respectively. For example, in the case of processive motor proteins a pairwise histogram directly yields the step height [2,11]. Further analysis of the variance among different displacement records can also provide estimates of the mean step size when individual steps cannot be recognized, but requires assumptions about the statistical distribution of step durations [12–16]. Furthermore, these methods require that the step height not vary within the data record under analysis.

Other methods based on the theory of hidden Markov models are quite successful, but also make use of assumed underlying kinetic models to locate steps of arbitrary size and directly estimate the parameters of the model [17,18]. When no such model is known or when assuming a model is undesirable, more general automatic methods, including velocity thresholding used by Hua et al. [19], wavelet component or multiscale product thresholding as discussed by Wang [20] or Sadler and Swami [21], a sliding

* Corresponding author. Tel.: +1 (520) 621 4276.
E-mail addresses: kalafut@physics.arizona.edu (B. Kalafut), visscher@physics.arizona.edu (K. Visscher).

0010-4655/$ – see front matter
doi:10.1016/j.cpc.2008.06.008

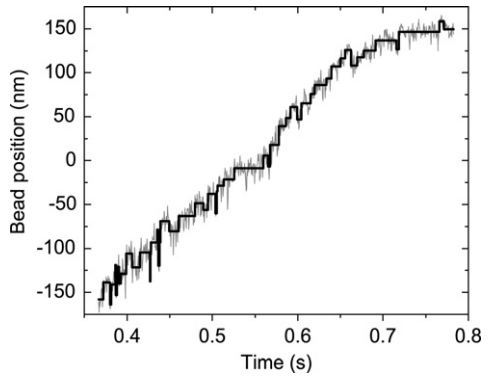

Fig. 1. Displacement record of a kinesin motor protein pulling a bead in a 5 pN force clamp at 5 μM ATP concentration [15] Grey line: Raw data trace. Black line: Result of our step-fitting method.

window Student t-test used by Carter and Cross [3], and step-function fitting with an heuristic termination criterion as developed by Kerssemakers [22] are available.

All of the above methods take from the user at least one parameter, usually corresponding to a physical time, length, or intensity scale. The accuracy with which steps are detected depends, often strongly, on the choice of this parameter, with a tradeoff between missing rapid or small steps and splitting single events into multiple steps or introducing spurious steps. For example, the sliding-window t-test yields ambiguous and, at times, unreliable results when two steps are near to each other in time. The step-fitting method of Kerssemakers et al. is notable in that its user-selected parameter corresponds not to a scale but to the number of steps fit to the data, with an heuristic, the ratio between χ^2 for the fit and χ^2 for a "counter-fit" with steps placed in the best fit's dwells, introduced to guide selection [22]. The authors note that the optimal parameter differs slightly from their theoretical framework's prediction, but provide no quantitative means to make the adjustment.

A recent review by Carter et al., comparing four model-independent step detection methods by Carter, Hua, Sadler, and Kerssemakers, found that the Kerssemakers method was the most accurate given the chosen criteria, especially when the signal-to-noise ratio approached 1.6, the lowest level tested [23]. Parameters of the other three methods were empirically optimized using simulated data; the Kerssemakers method's parameter was taken as its theoretical optimum. Pre-filters yielding the best result were likewise chosen empirically for each method using simulated data. No method was accurate enough at high noise levels to be compelling, moreover, such empirical parameter optimization and filter selection is not available to experimenters who cannot make assumptions about the underlying kinetic process.

We present a method for model-independent (i.e. not assuming a kinetic model) step detection which makes use of the Schwarz Information Criterion for statistical model selection to determine the number and location of steps [24]. Importantly, our method takes no input from the user other than the data itself, and makes no use of any assumed threshold values or confidence levels. Detection of steps is thus entirely automatic and objective, based solely on the properties of the data. To illustrate the challenges with which one is presented when identifying individual steps in typical single-molecule biophysical data, the resulting fit to the position of a single kinesin molecule pulling a bead against a constant applied force is plotted in Fig. 1 [15].

2. Objective test statistic for finding steps

Determining the number and location of steps in noisy data is an example of a general class of problem known as change-point problems, which involve estimating the number and location of changes in the underlying distribution of a series of random trials, most usually a change in statistical moments, such as the mean or variance [25]. Change-point problems have been extensively treated both as a problem of theoretical statistics and in applications to industrial quality control and finance [26,27].

For step detection we will test for mean change points only, as the strength of the underlying noise is assumed constant. However, the method can readily be extended to include variance change points or joint mean and variance changepoints; this may be used in the analysis of single-molecule data as obtained for, e.g., myosin II using optical tweezers, where changes in mean and variance are observed when myosin II interacts with actin [28–30].

Identification of mean change points requires discrimination between formal statistical models with varying numbers of change points or steps. Likelihood-ratio tests are in principle ill-suited to this task, as they tend always to favor models with higher numbers of parameters and thus the detection of as many steps as possible. Test statistics known as information criteria extend the likelihood-ratio procedure to allow discrimination between models with differing numbers of parameters by introducing rigorous, information-theoretically justified penalties for overfitting. Two of the more commonly used methods are the Akaike Information Criterion,

$$\mathrm{AIC} = -2\log\mathcal{L}(\hat{\theta}) + 2p,$$

which selects the model which results in the maximum information entropy, and the Schwarz Information Criterion,

$$\mathrm{SIC} = -2\log\mathcal{L}(\hat{\theta}) + p\log n, \tag{1}$$

which approximately chooses the model which yields the shortest possible encoding of the data and model [24,31–33]. In both cases, $\mathcal{L}$ is the likelihood function of the statistical model (see Appendix A), $\hat{\theta}$ is the vector of parameters (in our case, means and variances) which maximize the likelihood, p is the number of parameters, and n is the number of data in the set. The Schwarz Information Criterion is also known as the Bayesian Information Criterion or Schwarz–Bayes Criterion, as Schwarz derived it as a general approximation to a Bayes factor. We avoid such usage, as we do not use the SIC to estimate posterior probabilities or otherwise employ it in a Bayesian context. We choose to work with the SIC since, when biased, it tends to be biased towards underfitting, whereas the AIC tends to be biased towards overfitting [34]. In most biophysical contexts, we believe that it is better to miss small events than to introduce spurious ones.

We obtain (Appendix A) the SIC statistic for detecting mean changepoints by consider our data as a series of independent trials of Gaussian random variables with identical variances and directly substituting the Gaussian likelihood function into Eq. (1), yielding

$$\mathrm{SIC}(j_1,\ldots,j_k) = (k+2)\log n + n\log\hat{\sigma}^2_{j_1,\ldots,j_k} + n\log 2\pi + n. \tag{2}$$

$j_1,\ldots,j_k$ are the assumed step locations, and $\hat{\sigma}^2_{j_1,\ldots,j_k}$ is the maximum likelihood estimator of the variance computed assuming these step locations. The constant terms n and $n\log 2\pi$ may be dropped, as we are only interested in relative values of the SIC. In a series with k steps there are $k+1$ dwells and thus $k+1$ values of the mean, and one value for the global variance, giving $p = k+2$.

3. Successive step placement

Schwarz Information Criterion hypothesis tests are simple in principle: using all of the data, one computes a single number, the SIC, for each hypothesis in question and accepts the hypothesis for which this number is the lowest. As the hypotheses are judged only with respect to each other, there is no confidence level

or assumed threshold assumed or hidden in the algorithm. In the step-detection case, these competing hypotheses are the different numbers and locations of steps. Despite such conceptual simplicity, an SIC test of all change-point hypotheses becomes computationally costly very quickly; in the case of a data set of n entries one would in principle have to test $\binom{n}{k}$ configurations for the hypothesis that there are k steps, with 2^n possibilities overall.

A procedure known as "binary segmentation" allows this computationally costly problem to be reduced to testing for a single change point, then breaking the data set in two at the change point location, and repeating the test on the resulting subsequences until no more steps are found [27,35]. This reduces the step-finding problem to iteratively testing $n-1$ models containing a single mean change point, $\mathcal{H}_1(k)$, against a model containing no such change point, $\mathcal{H}_0$, with k the index of the location of the change point within the data set. Such a binary-segmenting method has been successfully applied to detection of intensity changepoints in series of time-resolved Poissonian photon counts [36].

However, we found that for our application binary segmentation introduces an unacceptable amount of overfitting, tending to break long dwells into pieces and to place several detected steps near a single stepping event. Moreover, the penalty for adding another step *decreases* at each iteration. Our attempts to remedy this by introducing a confidence level and an associated threshold value, as was done by Chen and Gupta for joint mean and variance changes, yielded only minor improvement [29]. Binary segmentation negates one of the key advantages of an information criterion test, as only the first step is placed and tested against the previous hypothesis using all of the data. Furthermore, the introduction of a user-tunable confidence level defeats the goal of full objectivity.

Instead of segmenting the data at a detected change point, we locate changepoints one at a time, using all of the data, as follows:

(1) For each possible location of a single step in the data, we calculate an SIC statistic, and accept the location for which the SIC is minimal as the best one-step fit.
(2) Then, again using the SIC, we test this best one-step hypothesis against the no-step hypothesis. If the no-step hypothesis is better, we terminate and conclude that there are no steps.
(3) Holding the location of the first step fixed, we compute an SIC for each possible assumed location of a second step, and determine the best such location. (See Section 5 for alternatives to fixed step placement.)
(4) We test this two-step hypothesis against the previous one-step hypothesis using the SIC, again terminating if the test is failed.
(5) Proceeding in a like fashion, we fit the kth step holding the previous $k-1$ steps fixed, and test the best k step hypothesis against the $k-1$ hypothesis, terminating and determining that there are $k-1$ steps when the when the $k-1$ step hypothesis has lower SIC than the best k step hypothesis.

Fig. 2 shows the result of the first two iterations and the final result of this algorithm's analysis of a sample data set.

The number of operations performed by this step-finding algorithm scales approximately linearly with the number of real steps present in the data and quadratically with the number of points in the set. In our implementation we reduce the number of multiplications per iteration by avoiding repeat calculation of variances.

4. Results

Applications of our step detection method are shown in Figs. 1 and 2, finding steps in a displacement of the motor protein kinesin and in simulated data traces, respectively. However, objective tests are required to quantitatively evaluate the performance of our method. To facilitate a direct comparison to existing methods, we have adopted the test criteria introduced by Carter et al., who recently reviewed the most commonly used step detection methods [23]:

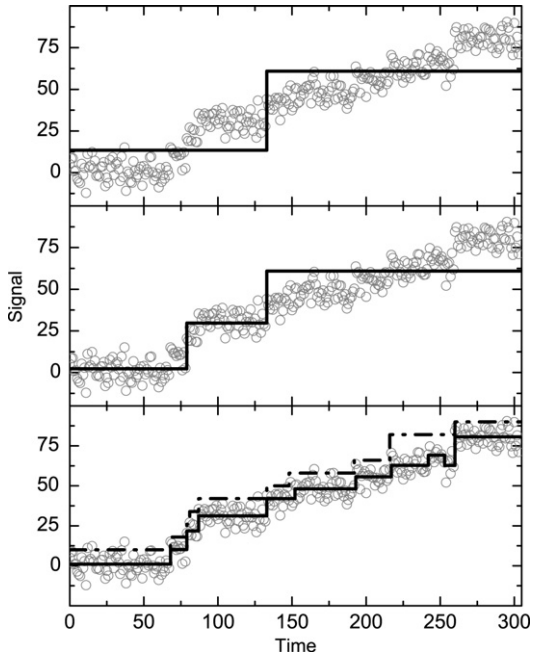

Fig. 2. Example of a fit to a simulated series of ten fixed-size steps, at $S/N = 8/5$. Dwell durations are governed by single-exponential kinetics, with a mean dwell time of 24 units. Grey circles are the raw, noisy simulated signal. Top: A single step is fit using the SIC statistic. Middle: A second step is added, holding the location of the previous step fixed. Bottom: The final result, after termination, with the underlying noiseless signal for reference (dashed line, offset for clarity).

(1) Successful identification of a particular real step is defined as the case in which the nearest detected step is no more than 2 data points away from the real step.
(2) A detected step is defined to correspond to a real step if the nearest real step is no more than two points away from the detected step.
(3) A step is of the correct height if it is within 3 units of the expected 8 unit step height.

However, since our method is not intrinsically windowed, counting the number of detected (real) steps for which there is a real (detected) step within a tolerance window, as in Carter's criteria 1 and 2, does not adequately characterize performance. The distribution of distances in time from real steps to the closest detected step, and detected steps to closest real steps, are better measures of the ability to find real steps and the quality of detected steps, respectively. Since Carter's criterion 3 allows for a 37.5% error in step heights, we also examine the actual distribution of recovered step heights. We first present an evaluation of our method based on study of these distributions, and then apply Carter's criteria for a direct comparison to existing methods.

Fig. 3 shows the distributions of distances in time from real steps to the closest detected steps (A, C) and from detected steps to the closest real steps (B, D) at signal-to-noise (S/N) ratios of 4 (A, B) and 2 (C, D). We define S/N as the ratio of the fixed step height to the standard deviation of the added Gaussian noise. 100 simulated time series were generated (Section 6), each with 200 steps and exponentially distributed dwell durations (mean = 24 data points). It is clear that all distributions peak sharply at a distance of 0 data points, indicating that steps are identified exactly practically 100% of the time at $S/N = 4$ to 90% of the time

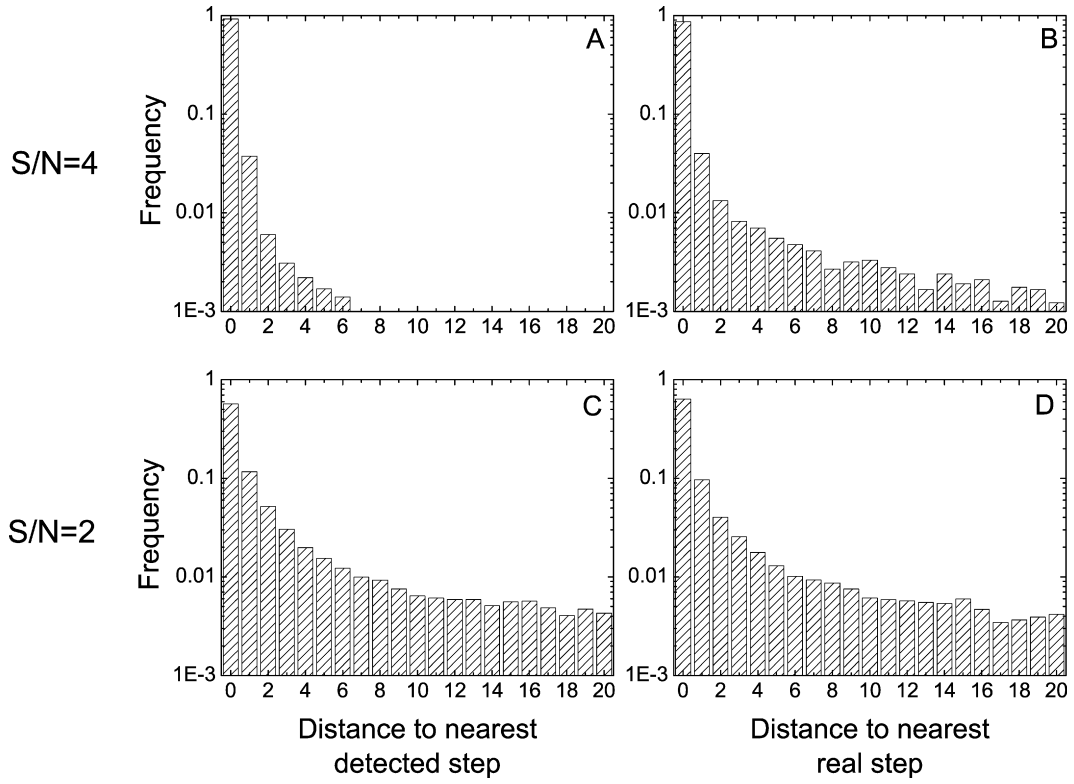

Fig. 3. Distributions of the distance in time, measured in data points, from real steps to the closest detected step and of the distance from detected steps to the closest real step, are plotted on a semilogarithmic scale at $S/N = 2$ and $S/N = 4$.

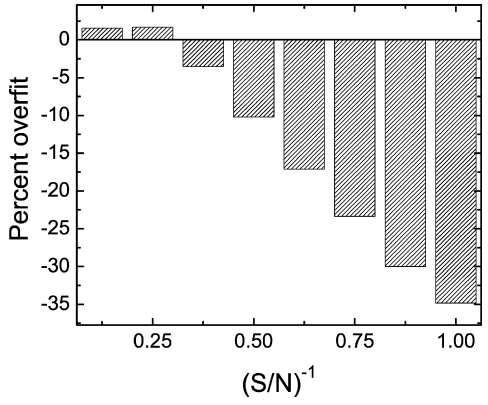

Fig. 4. Percent net overfit, defined as $100 \cdot (n_{\text{detected}} - n_{\text{real}})/n_{\text{real}}$, is plotted for the data series at eight different signal-to-noise ratios. Negative values correspond to underfitting.

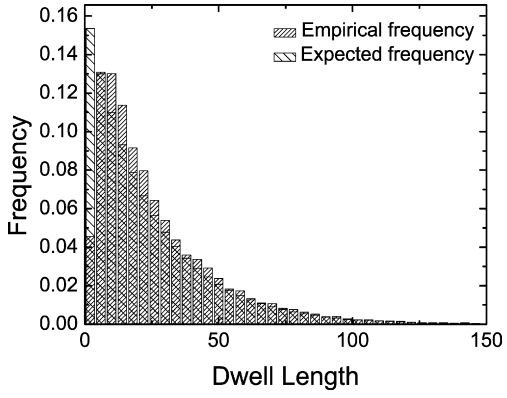

Fig. 5. The histogram of dwell lengths recovered by our method (stripes from lower left to upper right) is compared to the expected exponential distribution of dwell lengths (stripes from upper left to lower right) for $S/N = 2$.

at $S/N = 2$. Application of Carter's criterion, allowing a two-point window, increases successful detection of real steps and accurate placement of detected steps to nearly 100% at both noise levels, as can be seen from inspection of the histograms.

That, at $S/N = 4$, the distribution of distances from detected steps to real steps has a more extended tail than the distribution of distances from real steps to detected steps, reflects the tendency, as shown in Fig. 4, to slightly overfit at low noise levels. Underfitting at higher noise levels is characteristic of a well-behaved step-detection method; we miss small steps buried in or blurred together by the noise, rather than filling in noise runs at random. Note, however that the events in the tail are below the 1 percent level.

It is important to realize that even the slightest overfitting, underfitting, or incorrect positioning of steps adversely affects the determination of the distribution of dwell durations. In particular, we miss short dwells, corresponding to rapid stepping, as is illustrated in Fig. 5, which shows the empirical and expected distributions of detected dwell durations at $S/N = 2$. This biases the recovered mean dwell, in our case yielding 26.4 rather than 24, and illustrates once again that the possibility of missing evens should give rise to caution when interpreting (and fitting) such empirical distributions.

Fig. 6 shows the expected distribution of step heights and the empirical distribution recovered by application of our step-detection method to 100 simulated data series with 200 steps each with fixed 24-point dwell times and exponentially distributed step heights, with the mean height being twice the noise level (i.e. the S/N remains constant independent of step height). This distribution of step heights does not correspond to any biophysical process or experiment, but is useful in determining the step scales at which our step-detection method fails. We miss many steps

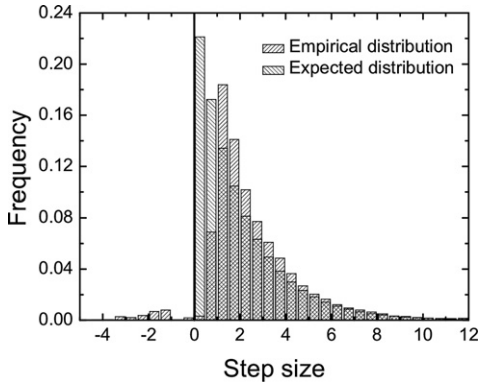

Fig. 6. We compare the histogram of recovered step heights (stripes from lower left to upper right) with the expected distribution (stripes from upper left to lower right) for a set of 100 data series, each with 200 evenly-spaced stepping events with exponentially distributed heights, with mean height equal to twice the noise level.

smaller than the noise level, and consequently find more steps slightly larger than the noise level than expected, once again biasing the mean rightward.

For direct comparison to existing methods as reviewed by Carter et al. we tested our method on data simulated as in the review (Section 6), consisting of a high-resolution series of fixed-height steps with exponentially distributed dwell lengths, smoothed by a binning process corresponding to instrumental time-averaging, to which Gaussian noise is subsequently added, with a 1.25/s mean stepping rate and 30 Hz sampling rate, allowing multiple steps to take place within the same frame. Each data set contains 200 stepping events; 100 trials were done at each noise level. Results are plotted in Fig. 7. At the highest noise level tested, $(S/N)^{-1} = 0.625$ (corresponding to 8 unit steps and a 5 unit noise standard deviation, adopting notation consistent with Carter et al.), we find approximately 5% more steps than the best reviewed method. Approximately 5% more of our detected steps are within two points of a real step than any of the reviewed methods without pre-filtering. Since our algorithm, by construction, makes use of a parametric model of the noise, we do not apply any pre-filter. Our performance in this test is equal to that of the best reviewed method applied to pre-filtered data. However, substantially more of our steps are of the expected height than any of the reviewed methods, with or without filters. The percentage of steps of the correct height *decreases* at the lowest noise level (Fig. 6, $(S/N)^{-1} = 0.125$); here our algorithm locates the one-point artifactual intermediate step created by the binning process.

5. Discussion

Our method takes no input from the user except the data itself. We do no internal heuristic parameter setting; the only assumption made is that the data take the form of steps with Gaussian white noise. In a sense, the algorithm presented above is, to the best of our knowledge, the first *fully objective* model-independent method for detection of steps in noisy data. We have found that under most circumstances, especially at high noise, we exceed the performance of step-finding algorithms with user-set parameters [23].

The major assumption inherent in our method is that of Gaussian noise. This is natural for experimental systems, such as optical tweezers, in which much of the noise arises from Brownian motion. Due to the Central Limit Theorem, assumption of Gaussian noise is also generally applicable to systems with several finite-variance noise sources. When the system is known to be governed by other underlying distributions, different test statistics, such as those derived by Watkins and Yang for the case of Poissonian photon counts, apply [36]. Likewise, variance estimation must be modified when the noise spectrum is not nearly white and the trials consequently not independent.

Our method of iteratively fitting steps, fixing the previous step locations, is one of many imaginable ways to practically implement an information criterion method for detection of multiple changepoints. For example, the photon-count changepoint algorithm of Watkins and Yang takes a different approach, using the result of binary segmentation as an input to a process which builds candidate state models (from a family based on assumptions about the underlying physics) and selects the best such model using a global SIC test [36]. For short data records, testing all $\binom{n}{k}$ configurations of k steps directly may itself be possible. When this is not a realistic option, further procedures may be added to our step-finding algorithm to increase accuracy at the expense of computational efficiency by allowing the changepoints to move at each iteration after placement. For example, one may, after placing each step, further optimize the fit by a combinatorial search of the neighborhoods around each step before testing for the existence and location of another step. The end result of our algorithm could likewise be used as the initial guess for a more general search of the parameter space. Given the performance of our algorithm as quantified above, we expect such computationally expensive methods to yield but marginal improvement.

In deriving our SIC statistic, we modeled the data as a series of discrete random trials with an *instantaneous* mean shift. When continuous data are sampled by time averaging, an intermediate level may result, even if the change in the mean of the underlying continuous process is nearly instantaneous. At low noise levels, (Fig. 8, $(S/N)^{-1} = 0.125$) our algorithm tends to find some of these intermediates, as they are real (albeit short) dwells in the discrete time series, and extreme outliers with respect to neighboring dwells, thus overfitting the data. Such overfitting at low noise levels does not occur in the analysis of non-binned simulated data in which the changepoints are truly discrete. A test statistic based on a formal step model explicitly accounting for the method of discretization may yield better performance at low noise levels, alternately, one may merge, *post hoc*, lone one-point dwells with their nearest neighbors, as such dwells, even when “real”, are indistinguishable from artifactual intermediates.

Low-frequency noise or deterministic drift are not uncommon in single-molecule experiments. To get a sense for our method’s robustness in the presence of drift, we tested its performance in finding steps in simulated data in which drift was naively modeled as deterministic and linear in time. Adding linear drift equal to 1% of the mean stepping rate to the data yields performance nearly indistinguishable from that presented in Fig. 6. When drift approaches 10% of the mean stepping rate, we see additional overfitting, especially at low noise, but retain reasonable accuracy (Fig. 8). It would appear from Fig. 8 as though drift actually leads to greater accuracy, but this is likely to be caused by the monotonic form of the drift and the absence of backward stepping in the underlying signal; constant linear drift in the same direction as stepping in some ways increases the prominence of each step. Regardless, the apparent advantage disappears when drift is increased much higher, and we do not expect it to generalize to real experimental drift. As a rule of thumb, the effect of drift is slight as long as, within the duration of a dwell, the magnitude of the drift is considerably less than the noise, as generally steps much smaller than the noise tend to be missed (Fig. 6). If the drift can be approximated by a known or fit function or modeled as, e.g., $1/f$ noise, one may in principle correct for it by substituting more appropriate maximum likelihood estimators into the SIC function; steps will then take the form of changepoints in a constant offset term, and the presence or absence of these changepoints will enable one to distinguish statistically between noise and steps.

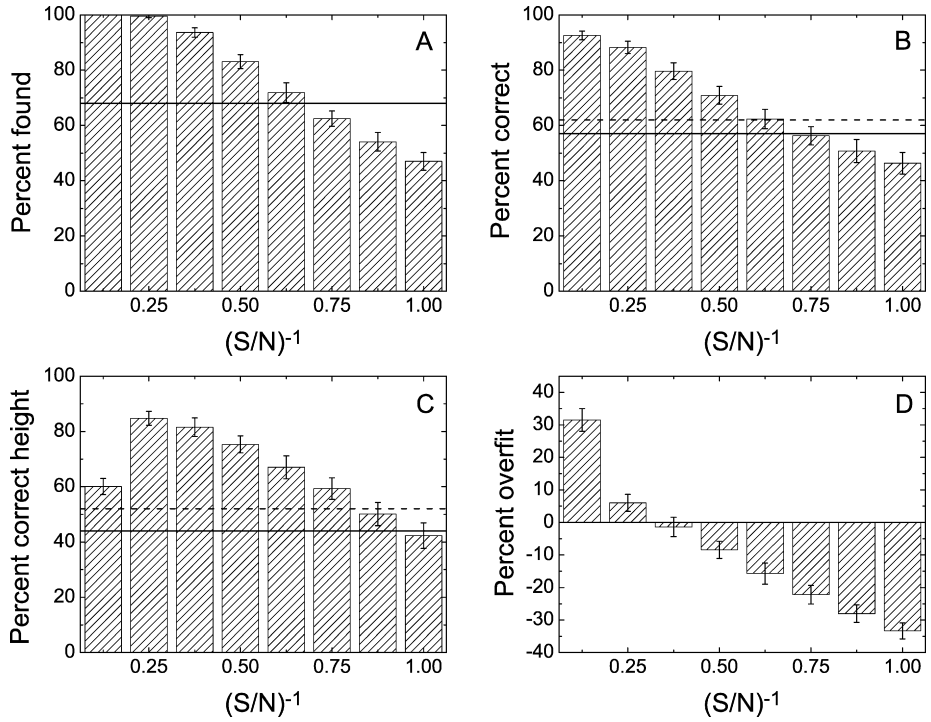

Fig. 7. To facilitate a direct comparison to other step detection methods, we perform the same tests of performance as Carter et al. [23] on data sets on which the action of a camera is simulated, and plot the results as a function of the noise level. Error bars represent standard deviations. The solid line represents the best performance at $(S/N)^{-1} = 0.625$ of any method tested by Carter without pre-filtering, and the dashed line represents best performance with pre-filtering, both estimated visually from figures contained in [23]. A: Ability to locate steps is quantified by counting the percentage of real steps for which there is at least one detected step placed within two sample points. Here the solid and dashed lines coincide. B: Quality of step placement is quantified by counting the percentage of detected steps for which there is at least one real step within two sample points. C: The tendency of our fitting routine to merge events or split single events into several sub-events is quantified by counting the percentage of detected steps with the expected 8 unit height (± 3). D: Overall tendency to over- or underfit is presented to clarify the implications of B and C.

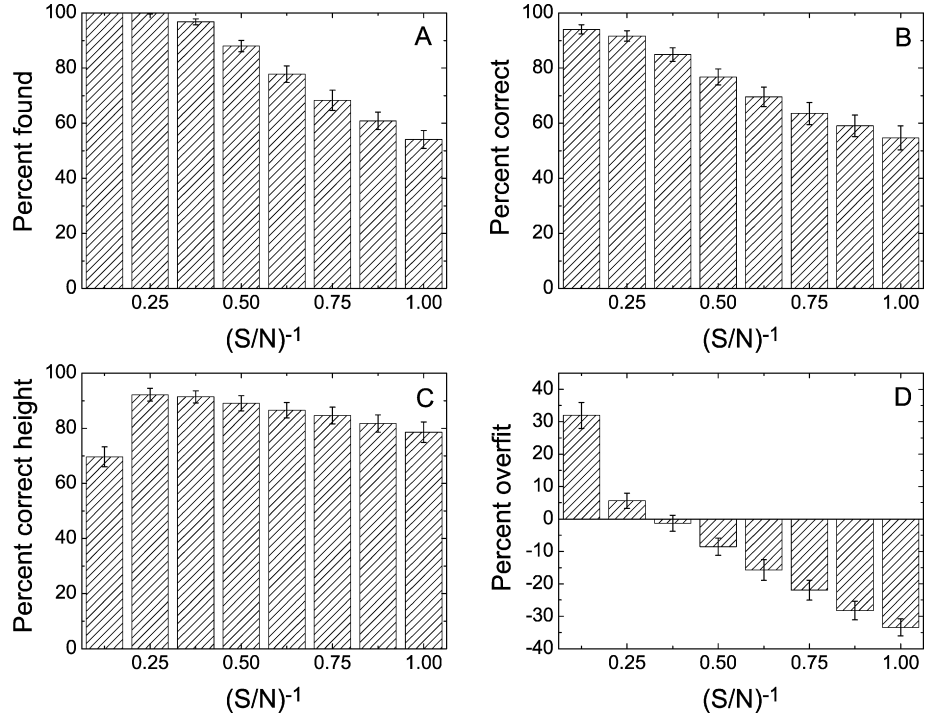

Fig. 8. Accuracy tests repeated on 100 data sets simulated as for Fig. 7, to which 10% linear drift has been added.

However, since in many cases such detailed characterization of drift is hard or even impossible, experimental designs using differential detection methods that subtract instrumental drift are much preferred. This has been successfully done for optical tweezers experiments by Shaevitz et al. [37] and by Nugent-Glandorf and Perkins [38] using a novel experimental geometry and fiducial marker bead, respectively. Absent experimental subtraction of drift, applicability of our step-detection method must be tested by

analysis of data simulated using a realistic model for noise and drift.

We have made the program code, compiled binaries, and LabView VIs available upon request by emailing

visscher@physics.arizona.edu.

6. Methods

6.1. Algorithm

Our step-finding algorithm was coded in C and interfaced to LabView (National Instruments, Austin, TX) using a Call Library Function node. A linked-list data structure was used to keep track of dwell durations, positions, and variances, reducing the number of multiplications to be performed per iteration.

6.2. Generation of benchmark data

To directly test our step-finding algorithm, we generate series of sharp steps with randomly distributed heights or dwell lengths (as specified above), adding Gaussian white noise. Unless otherwise noted, these series consist of 100 stepping events separated by exponentially distributed dwells with a mean length equal to 24 data points. "Noise level" refers to the standard deviation of the simulated Gaussian white noise.

In typical fluorescence or optical tweezers experiments, steps are not represented by sharp transitions in the data set; time averaging of positions or intensities smooths the transition slightly by inserting an artifactual intermediate state. Carter et al.'s review of four step detection methods tested their respective performance using such binned data as an input. To enable direct comparison, we followed their procedure in generating the benchmark data analyzed in Figs. 6 and 7. We first generate a 10 MHz series of dwells with exponentially distributed lengths. Single transition points are inserted in between the dwells to yield a transition time of 0.2 ms. This high-speed trace is split into segments which are averaged, simulating instrumental time-averaging. Finally, Gaussian white noise is added to this low-speed signal [23].

Generation of dwell times, step heights, and noise was done using LabView's built-in pseudorandom number generation.

Acknowledgements

This work was supported by the BIO5 Institute. B. K. was the recipient of a Biology, Math, and Physics Initiative fellowship.

We are grateful to Joe Watkins for discussions, Jie Chen for clarification of some of the mathematical literature, and Tom Perkins for testing several preliminary incarnations of our step-detection algorithm.

Appendix A. Derivation of the step-detection SIC

We recall that the formula for the normal distribution is

$$\mathcal{N}(\mu,\sigma^2)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{(x-\mu)^2}{2\hat{\sigma}^2}}$$

and that the maximum likelihood estimators for its mean and variance are

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

and

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2.$$

It follows immediately that the maximum likelihood estimator of the variance for a series of trials, assuming mean shifts (steps) at $j_1,\ldots,j_k$ and no variance shifts, where μ_i is the maximum likelihood estimator of the mean of the ith dwell and taking $j_0=0$ and $j_{k+1}=n+1$ for notational simplicity, is

$$\hat{\sigma}^2_{j_1,\ldots,j_k}=\frac{1}{n}\sum_{i=0}^{k}\sum_{l=j_i}^{j_{i+1}-1}(x_l-\mu_i)^2. \tag{3}$$

Using these estimators, the maximum likelihood function assuming these step locations is given by

$$\mathcal{L}(x_1,\ldots,x_n)=\prod_{i=0}^{k}\prod_{l=j_i}^{j_{i+1}-1}\frac{1}{\hat{\sigma}_{j_1,\ldots,j_k}\sqrt{2\pi}}e^{-\frac{(x_l-\mu_i)^2}{2\hat{\sigma}^2_{j_1,\ldots,j_k}}} \tag{4}$$

and its logarithm is thus

$$\log\mathcal{L}(x_1,\ldots,x_n)=-\frac{1}{2}\left(n\log 2\pi+n\log\hat{\sigma}^2_{j_1,\ldots,j_k}+\frac{1}{\hat{\sigma}^2_{j_1,\ldots,j_k}}\sum_{i=0}^{k}\sum_{l=j_i}^{j_{i+1}-1}(x_l-\mu_i)^2\right).$$

Recognizing the double sum in the last term to be $n\hat{\sigma}^2_{j_1,\ldots,j_k}$, we thus have

$$\log\mathcal{L}(x_1,\ldots,x_n)=-\frac{1}{2}\left(n\log 2\pi+n\log\hat{\sigma}^2_{j_1,\ldots,j_k}+n\right). \tag{5}$$

We substitute this into Eq. (1) and note that $p=k+2$ to obtain Eq. (2),

$$\mathrm{SIC}(j_1,\ldots,j_k)=(k+2)\log n+n\log\hat{\sigma}^2_{j_1,\ldots,j_k}+n\log 2\pi+n.$$

References

[1] E. Neher, B. Sakmann, Nature 260 (1976) 799.
[2] K. Svoboda, C.F. Schmidt, B.J. Schnapp, S.M. Block, Nature 365 (1993) 721.
[3] N. Carter, R. Cross, Nature 435 (2005) 308.
[4] S. Toba, T.M. Watanabe, L. Yamaguchi-Okimoto, Y.Y. Toyoshima, H. Higuchi, Proc. Natl. Acad. Sci. USA 103 (2006) 5741.
[5] T. Ha, et al., Proc. Natl. Acad. Sci. USA 96 (1999) 893.
[6] X. Zhuang, et al., Science 288 (2000) 2048.
[7] J.C.M. Gebhardt, A.E.M. Clemen, J. Jaud, M. Rief, Proc. Natl. Acad. Sci. USA 103 (2006) 8680.
[8] C. Kural, et al., Science 308 (2005) 1469.
[9] S. Chung, R. Kennedy, J. Neurosci. Methods 40 (1991) 71.
[10] M.J. de Castro, R.M. Fondecave, L.A. Clarke, C.F. Schmidt, R.J. Stewart, Nat. Cell. Biol. 2 (2000) 724.
[11] E.A. Abbondanzieri, W.J. Greenleaf, J.W. Shaevitz, R. Landick, S.M. Block, Nature 438 (2005) 460.
[12] K. Svoboda, S.M. Block, Cell 77 (1994) 773.
[13] K. Svoboda, P.P. Mitra, S.M. Block, Proc. Natl. Acad. Sci. USA 91 (1994) 11782.
[14] M.J. Schnitzer, S.M. Block, Nature 388 (1997) 386.
[15] K. Visscher, M.J. Schnitzer, S.M. Block, Nature 400 (1999) 184.
[16] K.C. Neuman, et al., Mat. J. Phys. Condens. 17 (2005) S3811.
[17] L.S. Milescu, A. Yildiz, P.R. Selvin, F. Sachs, Biophys. J. 91 (2006) 3135.
[18] L.S. Milescu, A. Yildiz, P.R. Selvin, F. Sachs, Biophys. J. 91 (2006) 1156.
[19] W. Hua, E.C. Yound, M.L. Fleming, J. Gelles, Nature 388 (1997) 390.
[20] Y. Wang, Biometrika 82 (1995) 385.
[21] B.M. Sadler, A. Swami, IEEE Trans. Inform. Theory 45 (1999) 1043.
[22] J. Kerssemakers, et al., Nature 442 (2006) 709.
[23] B.C. Carter, M. Vershinin, S.P. Gross, Biophys. J. 94 (2008) 306.
[24] G. Schwarz, Ann. Stat. 6 (1978) 461.
[25] E.S. Page, Biometrika 42 (1955) 523.
[26] G.A. Barnard, J. Roy. Stat. Soc. B 21 (1959) 239.
[27] J. Chen, A. Gupta, Commun. Statist. Simul. 30 (2001) 665.
[28] J. Chen, A. Gupta, J. Amer. Stat. Assoc. 92 (1997) 739.
[29] J. Chen, A. Gupta, Pap. Stat. 40 (1998) 323.
[30] C. Ruegg, et al., Sci. News Physiol. 17 (2002) 213.
[31] H. Akaike, IEEE Trans. Automat. Control. 19 (1974) 716.

[32] H. Akaike, On entropy maximization principle, in: P. Krishnaiah (Ed.), Applications of Statistics, North-Holland, 1977.
[33] J. Rissanen, Automatica 14 (1978) 465.
[34] K.P. Burnham, D.R. Anderson, Res. Soc. Meth. 33 (2004) 261.
[35] L.J. Vostrikova, Soviet Math. Dokl. 24 (1981) 55.
[36] L.P. Watkins, H. Yang, J. Phys. Chem. B 109 (2005) 617.
[37] J.W. Shaevitz, E.A. Abbondanzieri, R. Landick, S.M. Block, Nature 426 (2003) 684.
[38] L. Nugent-Glandorf, T.T. Perkins, Opt. Lett. 29 (2004) 2611.